

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
19 May 2005 (19.05.2005)

PCT

(10) International Publication Number
WO 2005/045835 A1

(51) International Patent Classification⁷: **G11B 20/10**

(21) International Application Number:
PCT/KR2004/002904

(22) International Filing Date:
10 November 2004 (10.11.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10-2003-0079181
10 November 2003 (10.11.2003) KR
10-2004-0083517 19 October 2004 (19.10.2004) KR

(71) Applicant: **SAMSUNG ELECTRONICS CO., LTD.**
[KR/KR]; 416, Maetan-dong, Yeongtong-gu, Suwon-si,
Gyeonggi-do 442-742 (KR).

(72) Inventors: **KANG, Man-Seok**; (301) 1237-3 Maetan
3-dong, Yeongtong-gu, Suwon-si, Gyeonggi-do 443-848
(KR). **MOON, Seong-Jin**; 403-506 Cheongmyung

Maeul 4-danji Apt., 1046-1, Yeongtong-dong, Yeong-
tong-gu Suwon-si, Gyeonggi-do 443-738 (KR). **CHUNG,**
Hyun-Kwon; 569-302 Shinsa-dong, Gangnam-gu, Seoul
135-891 (KR).

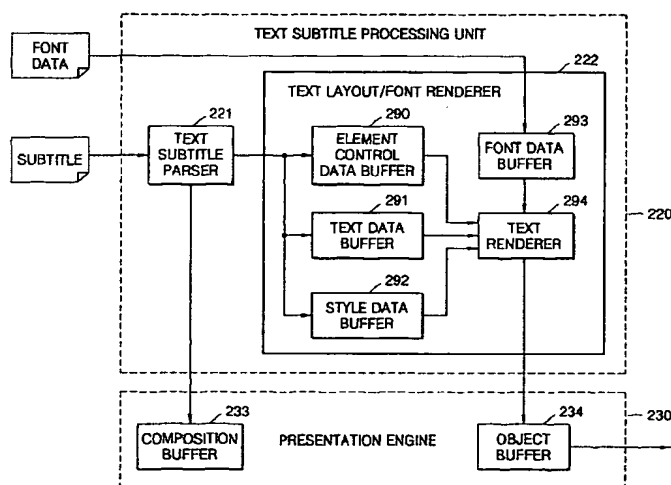
(74) Agent: **LEE, Young-Pil**; The Cheonghwa Building,
1571-18, Seocho-dong, Seocho-gu, Seoul 137-874 (KR).

(81) Designated States (*unless otherwise indicated, for every
kind of national protection available*): AE, AG, AL, AM,
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,
GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE,
KG, KP, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG,
MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH,
PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN,
TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every
kind of regional protection available*): ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,
FR, GB, GR, HU, IE, IS, IT, LU, MC, NL, PL, PT, RO, SE,
SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: INFORMATION STORAGE MEDIUM CONTAINING SUBTITLES AND PROCESSING APPARATUS THEREFOR



(57) Abstract: An information storage medium containing subtitles and a subtitle processing apparatus, where the information storage medium includes: audio-visual (AV) data; and subtitle data in which at least one subtitle text data and output style information designating an output form of the subtitle texts are stored with a text format. With this, output times of subtitle texts included in the text subtitle data can be overlapped, a subtitle file can be easily produced, and subtitles for an AV stream can be output with various forms.

WO 2005/045835 A1



Published:

— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Description

INFORMATION STORAGE MEDIUM CONTAINING SUBTITLES AND PROCESSING APPARATUS THEREFOR

Technical Field

- [1] The present invention relates to an information storage medium, and more particularly, to an information storage medium containing a plurality of subtitles that can be separately displayed and a processing apparatus therefor.

Background Art

- [2] A conventional subtitle is a bitmap image that is included in an audio-visual (AV) stream. Therefore, it is inconvenient to produce such a subtitle, and there is no choice but to merely read the subtitle in its present form without modification since a user cannot select various attributes of the subtitle defined by a subtitle producer. That is, since the attributes, such as font, character size, and character color, are predetermined and included in the AV stream as a bitmap image, the user cannot change the attributes at will.
- [3] Also, since the subtitle is compressed and encoded in the AV stream, an output start time and an output end time of the subtitle are clearly designated to correspond to the AV stream, and reproduction times when subtitles are output should not overlap. That is, only one subtitle should be output at a certain time.
- [4] However, since an output start time and an output end time of a subtitle are designated by a subtitle producer and recorded on an information storage medium separately from the AV stream, the output start times and output end times of a plurality of subtitles may overlap one another. In other words, since more than two subtitles may be output in a certain time period, a method of solving this problem is necessary.

Disclosure of Invention

Technical Solution

- [5] In an aspect of the present invention, the present invention provides an information storage medium having recorded thereon a plurality of text subtitles that are separately displayed although overlapping one another and an apparatus for reproducing the information storage medium.

Advantageous Effects

- [6] According to an embodiment of the present invention, a subtitle file can be easily produced, and subtitles for an AV stream can be output in various forms.

Description of Drawings

- [7] FIG. 1 illustrates a structure of a text subtitle file;
- [8] FIG. 2 is a block diagram of an apparatus reproducing an information storage medium on which a text subtitle is recorded;
- [9] FIG. 3 is a detailed block diagram of the text subtitle processing unit of FIG. 2;
- [10] FIG. 4 is a reference block diagram illustrating generation of a bitmap image without a presentation engine;
- [11] FIG. 5 is an example diagram illustrating correlations between structures in which composition information, position information, object information, and color information are recorded;
- [12] FIGS. 6A through 6C are diagrams illustrating a process of generating an image for a plurality of subtitles using one composition information data and one position information data;
- [13] FIGS. 7A through 7C are diagrams illustrating a process of generating an image for a plurality of subtitles using one composition information data and a plurality of position information data; and
- [14] FIGS. 8A through 8C are diagrams illustrating a process of generating an image so that one image object is included in one composition information data by allocating a plurality of composition information data for a plurality of subtitles.

Best Mode

- [15] According to an aspect of the present invention, there is provided an information storage medium including: AV data; and subtitle data in which at least one subtitle text data or output style information designating an output form of the subtitle texts are stored in a text format.
- [16] In an aspect of the present invention, the output style information contains pieces of information so that the output style information is differently applied to the subtitle texts.
- [17] In an aspect of the present invention, when a plurality of subtitle data exists, the plurality of subtitle data is separately rendered, and rendered images compose a plurality of pages, respectively.
- [18] According to another aspect of the present invention, there is provided a text subtitle processing apparatus including: a text subtitle parser separately extracting rendering information used to render a text from text subtitle data and control information used to present the rendered text; and a text layout/font renderer generating a bitmap image of a subtitle text by rendering the subtitle text according to the

extracted rendering information.

- [19] In an aspect of the present invention, the text layout/font renderer renders at least one subtitle text data by applying different styles to the subtitle text data and composes a plurality of pages with a plurality of rendered images.

Mode for Invention

- [20] Reference will now be made in detail to the embodiments of the present invention, examples of which are illustrated in the accompanying drawings, wherein like reference numerals refer to the like elements throughout. The embodiments are described below to explain the present invention by referring to the figures.

- [21] FIG. 1 illustrates a structure of a text subtitle file 100. Referring to FIG. 1, the text subtitle file 100 includes dialog information 110, presentation information 120, and meta data 130a and 130b.

- [22] The dialog information 110 includes subtitle texts, output start times of the subtitle texts, output end times of the subtitle texts, style groups or style information used to render the subtitle texts, text change effect information such as fade-in and fade-out, and a formatting code of the subtitle texts. The formatting code includes one or more of a code displaying a text with bold characters, a code for displaying the text in italics, a code indicating underlining, or a code indicating a line change.

- [23] The presentation information 120 includes style information used to render the subtitle texts and comprises a plurality of style groups. A style group is a bundle of styles on which the style information is recorded. A style includes information used to render and display a subtitle text. This information includes, for example, one or more of a style name, a font, a text color, a background color, a text size, a line-height, a text output region, a text output start position, an output direction, or an align method.

- [24] The meta data 130a and 130b, which are additional information of a moving picture, include information required to perform additional functions except a subtitle output function. For example, an additional function may be displaying a TV Parental Guideline such as 'TV-MA' on a screen for a program intended for mature audiences.

- [25] FIG. 2 is a block diagram of an apparatus reproducing an information storage medium on which a text subtitle file is recorded. It is understood that the apparatus may also record the text subtitle file to the information storage medium.

- [26] Referring to FIG. 2, a text subtitle processing unit 220 renders a subtitle text in order to process a text subtitle file. The text subtitle processing unit 220 includes a text subtitle parser 221, which extracts presentation information and dialog information from the text subtitle file, and a text layout/font renderer 222, which generates an

output image by rendering the subtitle text according to the extracted presentation information.

- [27] The text subtitle file 100 illustrated in FIG. 1 may be recorded on an information storage medium or in a memory included in a reproducing apparatus. In FIG. 2, the information storage medium or the memory on which the text subtitle file is recorded is called a subtitle information storage unit 200.
- [28] A text subtitle file corresponding to a reproducing moving picture and font data to be used to render the subtitle are read from the subtitle information storage unit 200 and stored in a buffer 210. The text subtitle file stored in the buffer 210 is transmitted to a text subtitle parser 221, which parses information required to render the text subtitle file. A subtitle text, font information, and rendering style information are transmitted to the text layout/font renderer 222, and control information of the subtitle text is transmitted to a composition buffer 233 of a presentation engine 230. The control information (i.e., information for displaying a screen with the subtitle text) includes an output region and an output start position.
- [29] The text layout/font renderer 222 generates a bitmap image by rendering the subtitle text using text rendering information transmitted from the text subtitle parser 221 and the font data transmitted from the buffer 210, composes one subtitle page by designating an output start time and an output end time of each subtitle text, and transmits the bitmap image and the subtitle page to an object buffer 234 of the presentation engine 230.
- [30] The subtitle of the bitmap image form read from the subtitle information storage unit 200 is input to a coded data buffer 231 and processed by a graphic processing unit 232 in the presentation engine 230. Accordingly, the graphic processing unit 232 generates a bitmap image. The generated bitmap image is transmitted to the object buffer 234, and control information of the bitmap image is transmitted to the composition buffer 233. The control information is used to designate a time and a position at which the bitmap image stored in the object buffer 234 is output to a graphic planer 240 and designating a color lookup table (CLUT) 250 in which color information to be applied to the bitmap image output to the graphic planer 240 is recorded. The composition buffer 233 receives object composition information transmitted from the text subtitle parser 221 and bitmap subtitle data processed by the graphic processing unit 232 and transmits control information for outputting the subtitle onto a screen to a graphic controller 235. The graphic controller 235 controls the object buffer 234 to combine the bitmap subtitle data processed by the graphic

processing unit 232 and rendered subtitle text object data received from the text layout/font renderer 222 and the graphic planer 240 to generate a graphic plane from the combined data, and outputs the graphic plane to a display unit (not shown) with reference to the CLUT 250.

[31] FIG. 3 is a detailed block diagram of the text subtitle processing unit 220 of FIG. 2. Referring to FIG. 3, a subtitle, which is text subtitle file information, is input to the text subtitle parser 221. The text subtitle parser 221 transmits subtitle control information parsed from the subtitle to the presentation engine 230 and text rendering information parsed from the subtitle to the text layout/font renderer 222. The text layout/font renderer 222 receives the text rendering information from the text subtitle parser 221 and stores control information of a subtitle text in an element control data buffer 290, subtitle text data in a text data buffer 291, and style information used to render the subtitle text data in a style data buffer 292. Also, the text layout/font renderer 222 stores font data used for text rendering in a font data buffer 293.

[32] The control information stored in the element control data buffer 290 may be a formatting code. The formatting code includes one or more of a code displaying a text with bold characters, a code displaying the text in Italics, a code indicating underlining, or a code indicating a line change. The subtitle text data stored in the text data buffer 291 is text data to be output as a subtitle. The style data stored in the style data buffer 292 may be one or more of data such as a font, a text color, a background color, a text size, a line-height, a text output region, a text output start position, an output direction, or an alignment method. A text renderer 294 generates a subtitle image with reference to the information recorded in each buffer and transmits the subtitle image to the presentation engine 230.

[33] FIG. 4 is a reference block diagram illustrating generation of a bitmap image without the presentation engine 230.

[34] That is, FIG. 4 illustrates another embodiment of an operation of the text subtitle processing unit 220 which includes a text subtitle controller 410 instead of the presentation engine 230.

[35] Referring to FIG. 4, the text layout/font renderer 222 generates composition information, position information, object information, and color information and generates a bitmap image on the basis of the composition information, the position information, the object information, and the color information. The text subtitle controller 410 receives object composition information from the text subtitle parser 221 and controls the text layout/font renderer 222 to directly output the bitmap image

generated by the text layout/font renderer 222 to the graphic planer 240 and the CLUT 250.

- [36] FIG. 5 is an example diagram illustrating correlations between structures in which composition information, position information, object information, and color information are recorded.
- [37] A subtitle that is output onto a screen is composed of page units. Each page can also include data used for other purposes besides the subtitle. The composition information refers to information containing information used to compose one page. The composition information includes output time information indicating a page output time, an object information reference value indicating an output image object, a position information reference value indicating an object output position, and a color information reference value indicating object color information.
- [38] The correlations between the information structures shown in FIG. 5 are a part of the composition information, and it is also possible to compose the correlations between position information, object information, and color information in a different form from FIG. 5.
- [39] Referring to FIG. 5, one page can include at least one region for outputting an image onto a screen. The at least one region is classified by the position information reference value. Position information refers to a recorded structure of information required to compose the at least one region to output the image. The position information includes information of horizontal and vertical coordinates of each region, a width of the region, and a height of the region. The object information includes object data to be displayed on the screen. Also, the object information includes object data type information corresponding to the object data.
- [40] An operation of the text subtitle processing unit 220 will be described as an example. The text subtitle processing unit 220 generates the composition information, the position information, the object information, and the color information of each rendered subtitle image to be output onto the screen in order to provide a subtitle text. The generated composition information, position information, object information, and color information are transmitted to the presentation engine 230.
- [41] As described above, when an information storage medium containing subtitles generated in a text form is reproduced, various exemplary methods of outputting more than one subtitle at the same time exist.
- [42] In a first method, the text subtitle processing unit 220 generates a new image for a plurality of subtitles, text output times of which are overlapped, and transmits a

subtitle composed of objects generated to be output to one position information in one composition information to the presentation engine 230.

[43] There is a second method of composing the subtitles so that the subtitles, text output times of which are overlapped, have different position information. That is, the text subtitle processing unit 220 generates an image of the plurality of subtitles, text output times of which are overlapped, using different position information data in one composition information and transmits the generated image to the presentation engine 230.

[44] There is a third method of generating subtitles, text output times of which are overlapped, using different composition information. That is, the text subtitle processing unit 220 generates different composition information data for a plurality of subtitles, text output times of which are overlapped, so that only one object is included in one composition information data.

[45] The three methods will be described in detail with reference to FIGS. 6 through 8.

[46] FIGS. 6A through 6C are diagrams illustrating a process of generating an image for a plurality of subtitles using one composition information data and one position information data.

[47] In FIG. 6A, a style 'Script' is defined as style information used for subtitle text rendering. Referring to FIG. 6A, the style 'Script' uses a font 'Arial.ttf', a text color 'black', a background color 'white', a character size '16pt', a text reference position of coordinates (x, y), an alignment method 'center', an output direction 'left-to-right-top-to-bottom', a text output region 'left, top, width, height', and a line-height '40px'.

[48] In FIG. 6B, subtitle texts 610, 620, and 630 rendered using the style 'Script' are defined. Referring to FIG. 6B, the subtitle text Hello 610 is output from '00:10:00' to '00:15:00', the subtitle text Subtitle 620 is output from '00:12:00' to '00:17:00', and the subtitle text World 630 is output from '00:14:00' to '00:19:00'. Therefore, two or three subtitle texts are output between '00:12:00' and '00:17:00'. Here, '
' indicates a line change. Using of the
 tag can prevent a plurality of subtitles from being overlapped on one region even though one style is used.

[49] FIG. 6C shows a result of outputting the subtitles defined in FIGS. 6A and 6B. Referring to FIG. 6C, data stored in each buffer of the text subtitle processing unit 220 in each illustrated time window will be described in detail.

[50] Before '00:10:00': the text layout/font renderer 222, when output composition information includes a void subtitle image, comprises:

- [51] Element control data buffer: void;
- [52] Text data buffer: void;
- [53] Style data buffer: style information of 'Script'; and
- [54] Font data buffer: font information of 'Arial.ttf'.
- [55] From '00:10:00' to '00:12:00': the text layout/font renderer 222, when output composition information including an image in which the subtitle text Hello 610 is rendered, comprises:
 - [56] Element control data buffer: control information of the subtitle text Hello 610;
 - [57] Text data buffer: 'Hello';
 - [58] Style data buffer: style information of 'Script'; and
 - [59] Font data buffer: font information of 'Arial.ttf'.
- [60] From '00:12:00' to '00:14:00': the text layout/font renderer 222, when output composition information including an image in which the subtitle text Hello 610 and the subtitle text Subtitle 620 are rendered, comprises:
 - [61] Element control data buffer: control information of the subtitle text Hello 610 and the subtitle text Subtitle 620;
 - [62] Text data buffer: 'Hello' and '
Subtitle';
 - [63] Style data buffer: style information of 'Script'; and
 - [64] Font data buffer: font information of 'Arial.ttf'.
- [65] From '00:14:00' to '00:15:00': the text layout/font renderer 222, when output composition information including an image in which the subtitle text Hello 610, the subtitle text Subtitle 620, and the subtitle text World 630 are rendered, comprises:
 - [66] Element control data buffer: control information of the subtitle text Hello 610, the subtitle text Subtitle 620, and the subtitle text World 630;
 - [67] Text data buffer: 'Hello' and '
Subtitle' and '

World';
 - [68] Style data buffer: style information of 'Script'; and
 - [69] Font data buffer: font information of 'Arial.ttf'.
- [70] From '00:15:00' to '00:17:00': the text layout/font renderer 222, when output composition information including an image in which the subtitle text Subtitle 620 and the subtitle text World 630 are rendered, comprises:
 - [71] Element control data buffer: control information of the subtitle text Subtitle 620 and the subtitle text World 630;
 - [72] Text data buffer: '
Subtitle' and '

World';
 - [73] Style data buffer: style information of 'Script'; and
 - [74] Font data buffer: font information of 'Arial.ttf'.

[75] From '00:17:00' to '00:19:00': the text layout/font renderer 222, when output composition information including an image in which the subtitle text World 630 is rendered, comprises:

[76] Element control data buffer: control information of the subtitle text World 630;

[77] Text data buffer: '

World';

[78] Style data buffer: style information of 'Script'; and

[79] Font data buffer: font information of 'Arial.ttf'.

[80] After '00:19:00': the text layout/font renderer 222, when output composition information includes a void subtitle image, comprises:

[81] Element control data buffer: void;

[82] Text data buffer: void;

[83] Style data buffer: style information of 'Script'; and

[84] Font data buffer: font information of 'Arial.ttf'.

[85] As shown in the above subtitle output process, in the first method, one subtitle image is generated by applying the same style to a plurality of subtitle texts having overlapped output times, one composition information data including the one subtitle image is generated, and the generated composition information data is transmitted to the presentation engine 230. At this time, page_time_out indicating the time when the transmitted composition information disappears from a screen refers to the time when a subtitle that is the last output to the screen among a plurality of subtitles having overlapped output times disappears or the time when a new subtitle is added.

[86] Text subtitle processing of the output subtitles must be quickly performed considering a time T_{decoding} taken to perform decoding of the subtitles in the text subtitle processing unit 220 and a time $T_{\text{composition}}$ taken to output the rendered subtitles from the object buffer 234 to the graphic planer 240. When T_{start} indicates the time when a subtitle is output from the text subtitle processing unit 220 of the reproducing apparatus, and when T_{arrival} indicates the time when the subtitle arrives at the text subtitle processing unit 220, correlations between these times are calculated by Equation 1.

[87]

Equation 1

$$T_{start} - T_{arrival} \geq T_{decoding} + T_{composition}$$

$$T_{decoding} = T_{rendering} + T_{composition \text{ information generation}}$$

$$T_{rendering} = \sum_{i=0}^{Num \text{ of } char} T_{char(i)}$$

[88] Referring to Equation 1, it can be known how quickly the text subtitle must be processed. Here, $T_{decoding}$ indicates the time taken to render a subtitle to be output, generating composition information including a rendered object, and transmitting the generated composition information to the object buffer 234. The subtitle requiring an output time of T_{start} must start to be processed before at least the time obtained by adding $T_{decoding}$ and $T_{composition}$. The time $T_{decoding}$ is obtained by adding $T_{rendering}$, which is the time taken to render the subtitle text and transmit the rendered subtitle text to the object buffer 234, and $T_{composition \text{ information generation}}$ which is the time taken to generate the composition information including the rendered object and transmit the composition information to the graphic planer 240. The time T_{char} is the time taken to render one character. Therefore, $T_{rendering}$ is obtained by adding times taken to render all characters.

[89] The size of the object buffer 234 must be equal to or larger than the size of the object. Here, the size of the object is obtained by adding the sizes of each character data of the object. Therefore, the number of characters composing one subtitle is limited to the number of characters which can be stored in the object buffer 234. Also, since the object buffer 234 can store a plurality of subtitles, the number of characters composing the plurality of subtitles is also limited to the number of characters which can be stored in the object buffer 234.

[90] FIGS. 7A through 7C are diagrams illustrating a process of generating an image for a plurality of subtitles using one composition information data and a plurality of position information data.

[91] In FIG. 7A, styles 'Script1', 'Script2', and 'Script3' are defined as style information used for subtitle text rendering. Referring to FIG. 7A, each of the three styles uses a font 'Arial.ttf', a text color 'black', a background color 'white', a character size '16pt', an

alignment method 'center', an output direction 'left-to-right-top-to-bottom', and a line-height '40px'. As a subtitle text reference position, 'Script1' has coordinates (x1, y1), 'Script2' has coordinates (x2, y2), and 'Script3' has coordinates (x3, y3). As a text output region, 'Script1' has 'left1, top1, width1, height1', 'Script2' has 'left2, top2, width2, height2', and 'Script3' has 'left3, top3, width3, height3'.

[92] In FIG. 7B, subtitle texts 710, 720, and 730 rendered using the styles 'Script1', 'Script2', and 'Script3' are defined. Referring to FIG. 7B, the subtitle text Hello 710 uses the style 'Script1' and is output from '00:10:00' to '00:15:00', the subtitle text Subtitle 720 uses the style 'Script2' and is output from '00:12:00' to '00:17:00', and the subtitle text World 730 uses the style 'Script3' and is output from '00:14:00' to '00:19:00'. Therefore, two or three subtitle texts are output between '00:12:00' and '00:17:00'. Since different scripts are used, the line change tag
 is unnecessary.

[93] FIG. 7C shows a result of outputting the subtitles defined in FIGS. 7A and 7B. Referring to FIG. 7C, data stored in each buffer of the text subtitle processing unit 220 in each illustrated time window will be described in detail.

[94] Before '00:10:00': the text layout/font renderer 222, when output composition information includes a void subtitle image, comprises:

[95] Element control data buffer: void;

[96] Text data buffer: void;

[97] Style data buffer: void; and

[98] Font data buffer: font information of 'Arial.ttf'.

[99] From '00:10:00' to '00:12:00': the text layout/font renderer 222, when output composition information includes an image in which the subtitle text Hello 710 is rendered, comprises:

[100] Element control data buffer: control information of the subtitle text Hello 710;

[101] Text data buffer: 'Hello';

[102] Style data buffer: style information of 'Script1'; and

[103] Font data buffer: font information of 'Arial.ttf'.

[104] From '00:12:00' to '00:14:00': the text layout/font renderer 222, when output composition information includes the subtitle text Hello 710 and the subtitle text Subtitle 720, comprises:

[105] Element control data buffer: control information of the subtitle text Hello 710 and the subtitle text Subtitle 720;

[106] Text data buffer: 'Hello' and 'Subtitle';

[107] Style data buffer: style information of 'Script1' and 'Script2'; and

- [108] Font data buffer: font information of 'Arial.ttf'.
- [109] From '00:14:00' to '00:15:00': the text layout/font renderer 222, when output composition information includes the subtitle text Hello 710, the subtitle text Subtitle 720, and the subtitle text World 730, comprises:
 - [110] Element control data buffer: control information of the subtitle text Hello 710, the subtitle text Subtitle 720, and the subtitle text World 730;
 - [111] Text data buffer: 'Hello', 'Subtitle', and 'World';
 - [112] Style data buffer: style information of 'Script1', 'Script2', and 'Script3'; and
 - [113] Font data buffer: font information of 'Arial.ttf'.
- [114] From '00:15:00' to '00:17:00': the text layout/font renderer 222, when output composition information includes the subtitle text Subtitle 720 and the subtitle text World 730, comprises:
 - [115] Element control data buffer: control information of the subtitle text Subtitle 720 and the subtitle text World 730;
 - [116] Text data buffer: 'Subtitle' and 'World';
 - [117] Style data buffer: style information of 'Script2' and 'Script3'; and
 - [118] Font data buffer: font information of 'Arial.ttf'.
- [119] From '00:17:00' to '00:19:00': the text layout/font renderer 222, when output composition information includes the subtitle text World 730, comprises:
 - [120] Element control data buffer: control information of the subtitle text World 730;
 - [121] Text data buffer: 'World';
 - [122] Style data buffer: style information of 'Script3'; and
 - [123] Font data buffer: font information of 'Arial.ttf'.
- [124] After '00:19:00': the text layout/font renderer 222, when output composition information includes a void subtitle image, comprises:
 - [125] Element control data buffer: void;
 - [126] Text data buffer: void;
 - [127] Style data buffer: void; and
 - [128] Font data buffer: font information of 'Arial.ttf'.
- [129] In the second method described above, subtitle images for subtitle texts are generated by applying different styles to a plurality of subtitle texts having overlapped output times, one composition information data including the subtitle images is generated, and the generated composition information data is transmitted to the presentation engine 230. A text subtitle processing time is the same as that of the first method. That is, text subtitle processing of the output subtitles must be quickly

performed considering a time T_{decoding} taken to perform decoding of the subtitles in the text subtitle processing unit 220 and a time $T_{\text{composition}}$ taken to output the rendered subtitles from the object buffer 234 to the graphic planer 240. However, in this method, since a plurality of objects exist, a rendering time is obtained by adding the times taken to render the respective objects. That is, Equation 2 calculates the rendering time.

[130]

Equation 2

$$T_{\text{start}} - T_{\text{arrival}} \geq T_{\text{decoding}} + T_{\text{composition}}$$

$$T_{\text{decoding}} = T_{\text{rendering}} + T_{\text{composition information generation}}$$

$$T_{\text{rendering}} = \sum_{i=0}^{\text{Num of obj}} T_{\text{OBJ}(i)}$$

$$T_{\text{OBJ}} = \sum_{i=0}^{\text{Num of char}} T_{\text{char}(i)}$$

[131] The number of characters of the subtitle text which can be stored in the object buffer 234 is limited in the second method to the same as that of the first method.

[132] FIGS. 8A through 8C are diagrams illustrating a process of generating an image so that one image object is included in one composition information data by allocating a plurality of composition information data for a plurality of subtitles.

[133] In FIG. 8A, styles 'Script1', 'Script2', and 'Script3' are defined as style information used for subtitle text rendering. Referring to FIG. 8A, each of the three styles uses a font 'Arial.ttf', a text color 'black', a background color 'white', a character size '16pt', an align method 'center', an output direction 'left-to-right-top-to-bottom', and a line-height '40px'. As a subtitle text reference position, 'Script1' has coordinates (x1, y1), 'Script2' has coordinates (x2, y2), and 'Script3' has coordinates (x3, y3). As a text output region, 'Script1' has 'left1, top1, width1, height1', 'Script2' has 'left2, top2, width2, height2', and 'Script3' has 'left3, top3, width3, height3'.

[134] In FIG. 8B, subtitle texts 810, 820, and 830 rendered using the styles 'Script1', 'Script2', and 'Script3' are defined. Referring to FIG. 8B, the subtitle text Hello 810 uses the style 'Script1' and is output from '00:10:00' to '00:15:00', the subtitle text

Subtitle 820 uses the style 'Script2' and is output from '00:12:00' to '00:17:00', and the subtitle text World 830 uses the style 'Script3' and is output from '00:14:00' to '00:19:00'. Therefore, two or three subtitle texts are overlapped between '00:12:00' and '00:17:00'.

- [135] FIG. 8C shows a result of outputting the subtitles defined in FIGS. 8A and 8B. Referring to FIG. 8C, data stored in each buffer of the text subtitle processing unit 220 in each illustrated time window will be described in detail.
- [136] From '00:00:00': the text layout/font renderer 222, when output composition information includes a void subtitle image, comprises:
- [137] Element control data buffer: void;
- [138] Text data buffer: void;
- [139] Style data buffer: void; and
- [140] Font data buffer: font information of 'Arial.ttf'.
- [141] From '00:10:00': the text layout/font renderer 222, when output composition information including an image in which the subtitle text Hello 810 is rendered, comprises:
- [142] Element control data buffer: load control information of the subtitle text Hello 810;
- [143] Text data buffer: 'Hello';
- [144] Style data buffer: style information of 'Script1'; and
- [145] Font data buffer: font information of 'Arial.ttf'.
- [146] From '00:12:00': the text layout/font renderer 222, when output composition information including the subtitle text Hello 810 and composition information including the subtitle text Subtitle 820 is rendered, comprises:
- [147] Element control data buffer: load control information of the subtitle text Subtitle 820;
- [148] Text data buffer: 'Subtitle';
- [149] Style data buffer: style information of 'Script2'; and
- [150] Font data buffer: font information of 'Arial.ttf'.
- [151] From '00:14:00': the text layout/font renderer 222, when output composition information including the subtitle text Hello 810, composition information including the subtitle text Subtitle 820, and composition information including the subtitle text World 830 is rendered, comprises:
- [152] Element control data buffer: load control information of the subtitle text World 830;
- [153] Text data buffer: 'World';
- [154] Style data buffer: style information of 'Script3'; and

- [155] Font data buffer: font information of 'Arial.ttf'.
- [156] After '00:15:00': the text subtitle processing unit 220 does not execute any operation until preparing an output for subsequent subtitle texts to be output after '00:19:00'. Therefore, changes of subtitles output between '00:15:00' and '00:19:00' are performed by the presentation engine 230 controlling the composition information of the subtitles 'Hello', 'Subtitle', and 'World' received from the text subtitle processing unit 220.
- [157] That is, at '00:15:00', the presentation engine 230 deletes the composition information and bitmap image object of the subtitle 'Hello' from the composition buffer 233 and the object buffer 234 and outputs only the composition information of the subtitles 'Subtitle' and 'World' onto a screen. At '00:17:00', the presentation engine 230 deletes the composition information and bitmap image object of the subtitle 'Subtitle' from the composition buffer 233 and the object buffer 234 and outputs only the composition information of the subtitle 'World' onto the screen. Also, at '00:19:00', the presentation engine 230 deletes the composition information and bitmap image object of the subtitle 'World' from the composition buffer 233 and the object buffer 234 and does not output a subtitle onto the screen any more.
- [158] In the third method described above, one subtitle image for each subtitle text is generated by applying different styles to a plurality of subtitle texts having overlapped output times, one composition information data is generated for each subtitle image, and the generated plurality of composition information data are transmitted to the presentation engine 230. A text subtitle processing time is the same as that of the first method. While only a processing time of only one composition information data is considered in the first and second methods since one composition information data for a plurality of subtitle texts having overlapped output times is composed and output, a plurality of composition information data are generated and output in the third method since each subtitle text composes a separate composition information data. Therefore, for a subtitle text processing start time of the third method, the worst case, that is, a case where a plurality of composition information data for a plurality of subtitles having the same output start time are simultaneously generated and output, must be considered. This is described by Equation 3.

[159]

Equation 3

$$\begin{aligned}
 T_{start} - T_{arrival} &\geq T_{decoding} + T_{composition} \\
 T_{decoding} &= T_{rendering} + T_{composition \text{ information generation}} \\
 T_{composition \text{ information generation}} &= \sum_{i=0}^{\text{number of composition information data}} T_{composition \text{ information}(i)} \\
 T_{rendering} &= \sum_{i=0}^{\text{Num of obj}} T_{OBJ(i)} \\
 T_{OBJ} &= \sum_{i=0}^{\text{Num of char}} T_{char(i)}
 \end{aligned}$$

[160] The time $T_{\text{composition information generation}}$ taken to generate a plurality of composition information data is obtained by adding each $T_{\text{composition information}}$ which is a composition information generation time of one subtitle, all together. The time $T_{\text{rendering}}$ taken to generate a plurality of objects by rendering a plurality of subtitles is obtained by adding each T_{OBJ} , which is a rendering time of one subtitle, all together. The time T_{OBJ} taken to render one subtitle is obtained by adding each T_{char} , which is a rendering time of each character included in a relative subtitle, all together. Referring to Equation 3, in order to simultaneously output a plurality of subtitles including a plurality of characters, a sum of times taken to render all characters included in the subtitles, compose the plurality of composition information data, and output the plurality of composition information data must be less than a difference between a subtitle output time and a subtitle processing start time of the text subtitle processing unit 220.

[161] The number of characters of the subtitle text that can be stored in the object buffer 234 is limited in the third method to the same as that of the first method or the second method.

[162] As described in the third method, in an information storage medium and a reproducing apparatus constructed with a structure supporting simultaneous output of a plurality of composition information data, a text subtitle and another bitmap image can be simultaneously output onto a screen.

[163] Data compressed and encoded in an AV stream includes video data, audio data, bitmap-based subtitles, and other non-subtitle bitmap images. An image 'TV-14' displayed on a top-right of a screen in order to indicate a TV program for people over

14 years old is an example of the non-subtitle bitmap images. In a conventional method, since only one composition information data is output onto a screen at one time, a region for outputting a bitmap subtitle and a region for outputting a non-subtitle bitmap image are separately defined in composition information in order to simultaneously output the bitmap subtitle and the non-subtitle bitmap image.

[164] Accordingly, when a user turns an output of subtitles off since the user does not want the output of the subtitles, a decoder stops only decoding of the subtitles. Therefore, since subtitle data is not transmitted to an object buffer, the subtitles disappear from a screen, and only the non-subtitle bitmap image is continuously output onto the screen.

[165] When the text subtitle processing unit 220 generates an image for a subtitle using one composition information data and transmits the composition information data to the presentation engine 230 in order to output the subtitle, if an output of subtitles is turned off, a non-subtitle bitmap image recorded in an AV stream is also not output. Therefore, in a case where a plurality of composition information data can be simultaneously output onto a screen as described in the third method of the present invention, when text subtitles are selected instead of bitmap subtitles, images except the bitmap subtitles in composition information included in an AV stream can be continuously output, and the text subtitles can be output using composition information generated by the text subtitle processing unit 220. That is, the text subtitles and the other non-subtitle bitmap images can be simultaneously output onto the screen.

[166] The present invention may be embodied in a general-purpose computer by running a program from a computer-readable medium, including but not limited to storage media such as magnetic storage media (ROMs, RAMs, floppy disks, magnetic tapes, etc.), optically readable media (CD-ROMs, DVDs, etc.), and carrier waves (transmission over the internet). The present invention may be embodied as a computer-readable medium having a computer-readable program code unit embodied therein for causing a number of computer systems connected via a network to effect distributed processing. And the functional programs, codes and code segments for embodying the present invention may be easily deducted by programmers in the art which the present invention belongs to.

[167] Although a few embodiments of the present invention have been shown and described, it would be appreciated by those skilled in the art that changes may be made in these embodiments without departing from the principles and spirit of the invention, the scope of which is defined in the claims and their equivalents.

Claims

- [1] 1. An information storage medium, comprising:
audio-visual (AV) data; and
subtitle data in which at least one subtitle text data and output style information designating an output form of the at least one subtitle text data are stored in a text format.
- [2] 2. The information storage medium of claim 1, wherein the output style information contains a plurality of pieces of information so that the output style information is differently applied to each subtitle text.
- [3] 3. The information storage medium of claim 1, wherein each subtitle text data is rendered by a reproducing apparatus applying a same output style according to the output style information and generating one page composed of one image corresponding to the subtitle data.
- [4] 4. The information storage medium of claim 1, wherein each subtitle text data is rendered by a reproducing apparatus applying different output styles and generating pages, each page comprising each corresponding rendered subtitle text data.
- [5] 5. The information storage medium of claim 1, wherein, when a plurality of subtitle data exist, the plurality of subtitle data are rendered as images, and the rendered images comprise a plurality of pages, respectively.
- [6] 6. The information storage medium of claim 1, wherein the subtitle data further comprises information of time when the at least one subtitle text is output onto a screen.
- [7] 7. A text subtitle processing apparatus, comprising:
a text subtitle parser separately extracting from subtitle data rendering information used to render a text in text subtitle data and control information used to present the rendered text; and
a text layout/font renderer generating a bitmap image of a subtitle text extracted by the text subtitle parser by rendering the subtitle text according to the extracted rendering information.
- [8] 8. The apparatus of claim 7, wherein the text subtitle parser extracts the control information so that the control information is fitted to a predetermined information structure format and transmits the control information to a presentation engine.

- [9] 9. The apparatus of claim 7, further comprising:
a text subtitle controller controlling the bitmap image generated by the text layout/font renderer to be directly output onto a screen using the control information separately from a presentation engine processing bitmap subtitle data.
- [10] 10. The apparatus of claim 7, wherein the subtitle data is a plurality of subtitle data having overlapping output times.
- [11] 11. The apparatus of claim 7, wherein the text layout/font renderer generates the bitmap image by generating one composition information data, one position information data, and one object information data corresponding to a plurality of the subtitle data having overlapping output times and outputs the bitmap image.
- [12] 12. The apparatus of claim 7, wherein the text layout/font renderer generates the bitmap image by generating one composition information data, one position information data, and a plurality of object information data corresponding to a plurality of the subtitle data having overlapping output times and outputs the bitmap image.
- [13] 13. The apparatus of claim 7, wherein the text layout/font renderer generates the bitmap image by generating a plurality of composition information data and one position information data and one object information data corresponding to each composition information data corresponding to a plurality of the subtitle data having overlapping output times and outputs the bitmap image.
- [14] 14. The apparatus of claim 7, wherein the text layout/font renderer renders one image of a plurality of the text subtitle data by applying a same output style to the plurality of the text subtitle data and generates one page comprising the one image.
- [15] 15. The apparatus of claim 7, wherein the text layout/font renderer renders images of a plurality of the text subtitle data by applying different output styles to each of the plurality of the text subtitle data and generates one page comprising a plurality of the rendered images.
- [16] 16. The apparatus of claim 7, wherein the text layout/font renderer renders images of a plurality of the text subtitle data by applying different output styles to each of the plurality of the text subtitle data and generates a plurality of pages comprising a plurality of the rendered images.
- [17] 17. A reproducing apparatus using an information storage medium, comprising:
a reader which reads a plurality of subtitles, each subtitle comprising subtitle texts, control information and style information from the information storage

medium ;

a buffer which buffers the subtitles, the subtitle texts, the control information, and the style information; and

a reproducer which decodes the subtitles based on the control information and the style information and displays the subtitle texts according to the control information and style information, wherein any number of the plurality of subtitles are simultaneously displayed.

- [18] 18. The apparatus of claim 17, wherein the reproducer comprises:
a subtitle processor which extracts the subtitle texts, the control information, and the style information from each of the subtitles and generates an output image by rendering the extracted subtitle texts according to the extracted control information and style information; and
a presentation engine which displays the output image.
- [19] 19. The apparatus of claim 18, wherein the reader reads a non-subtitle bitmap image and the presentation engine displays the non-subtitle bitmap image simultaneously with the output image.
- [20] 20. The apparatus of claim 17, wherein the reproducer comprises:
a subtitle processor which extracts the subtitle texts, the control information, and the style information from each of the subtitles and generates an output image by rendering the extracted subtitle texts according to the extracted control information and style information; and
a text subtitle controller controlling the output image generated by the subtitle processor to display on a screen according to the extracted control information and style information.
- [21] 21. A method of reproducing text subtitle files, comprising:
selecting subtitles having overlapping reproduction times for reproduction;
generating a single set of composition information data, position information data and object information data for the subtitles; and
generating an image corresponding to the subtitles according to the single set of composition information data, position information data and object information data.
- [22] 22. A method of reproducing text subtitle files, comprising:
selecting subtitles having overlapping reproduction times for reproduction;
generating a single set of composition information data having different position information data for each of the selected ones of the subtitles and different object

information data for each of the selected ones of the subtitles; and
generating an image corresponding to the subtitles according to the single set of
composition information data, the different position information data for each of
the subtitles and the different object information data for each of the subtitles.

- [23] 23. A method of reproducing text subtitle files, comprising:
selecting subtitles having overlapping reproduction times for reproduction;
generating a different set of composition information data, position information
data and object information data for each of the subtitles; and
generating an image corresponding to the subtitles according to the different sets
of composition information data, position information data and object in-
formation data for each of the subtitles.

FIG. 1

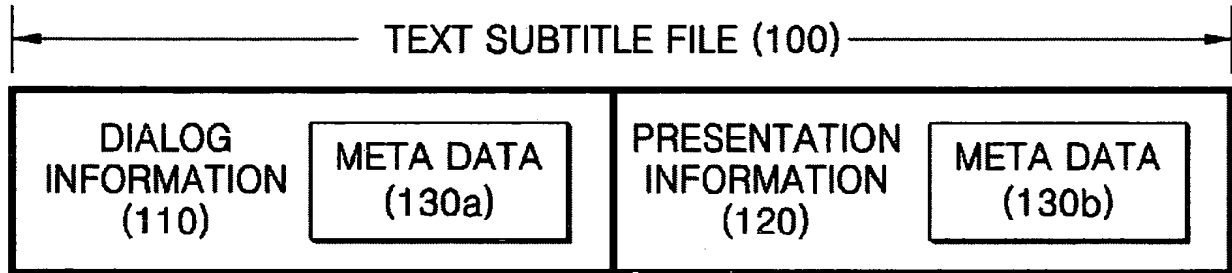


FIG. 2

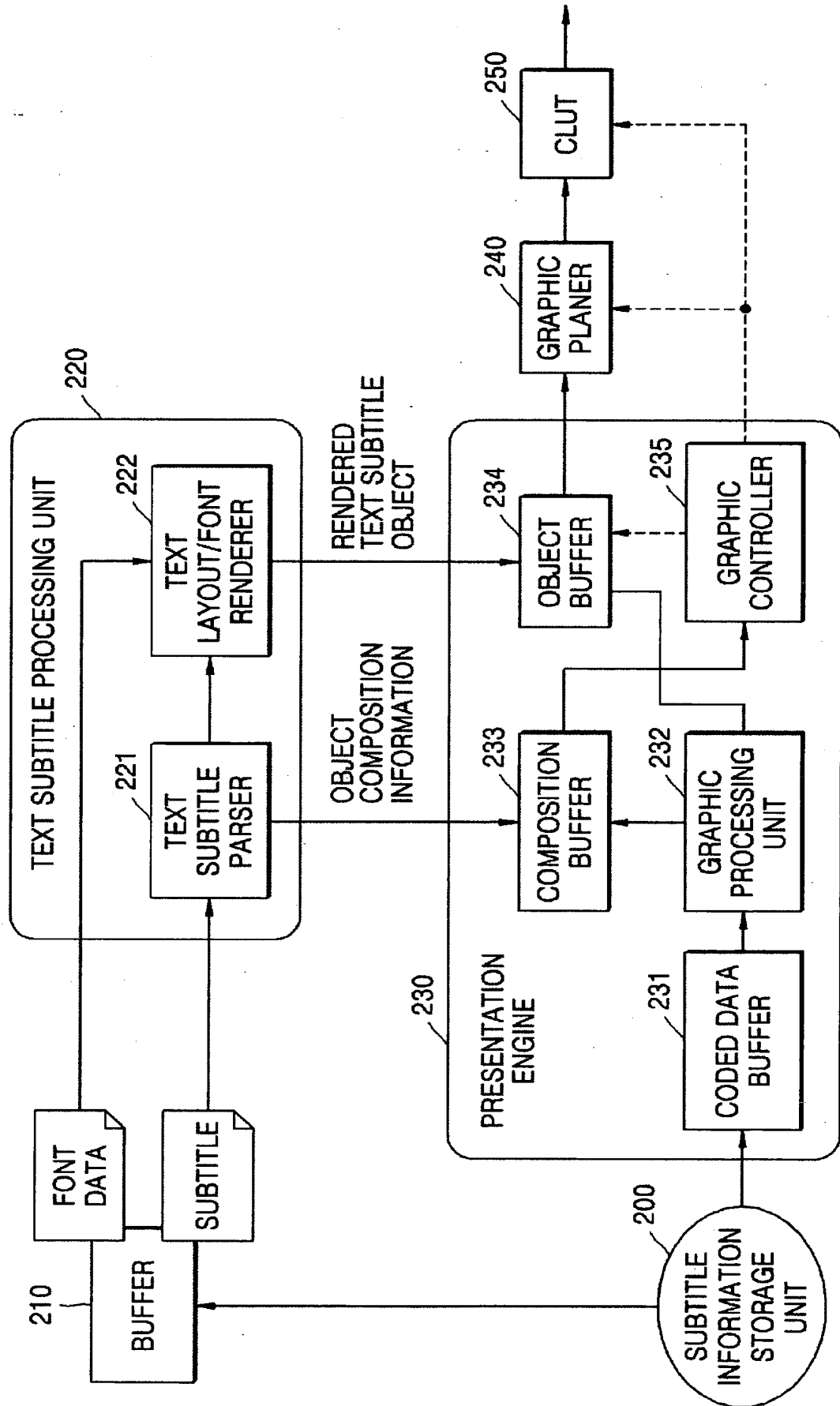


FIG. 3

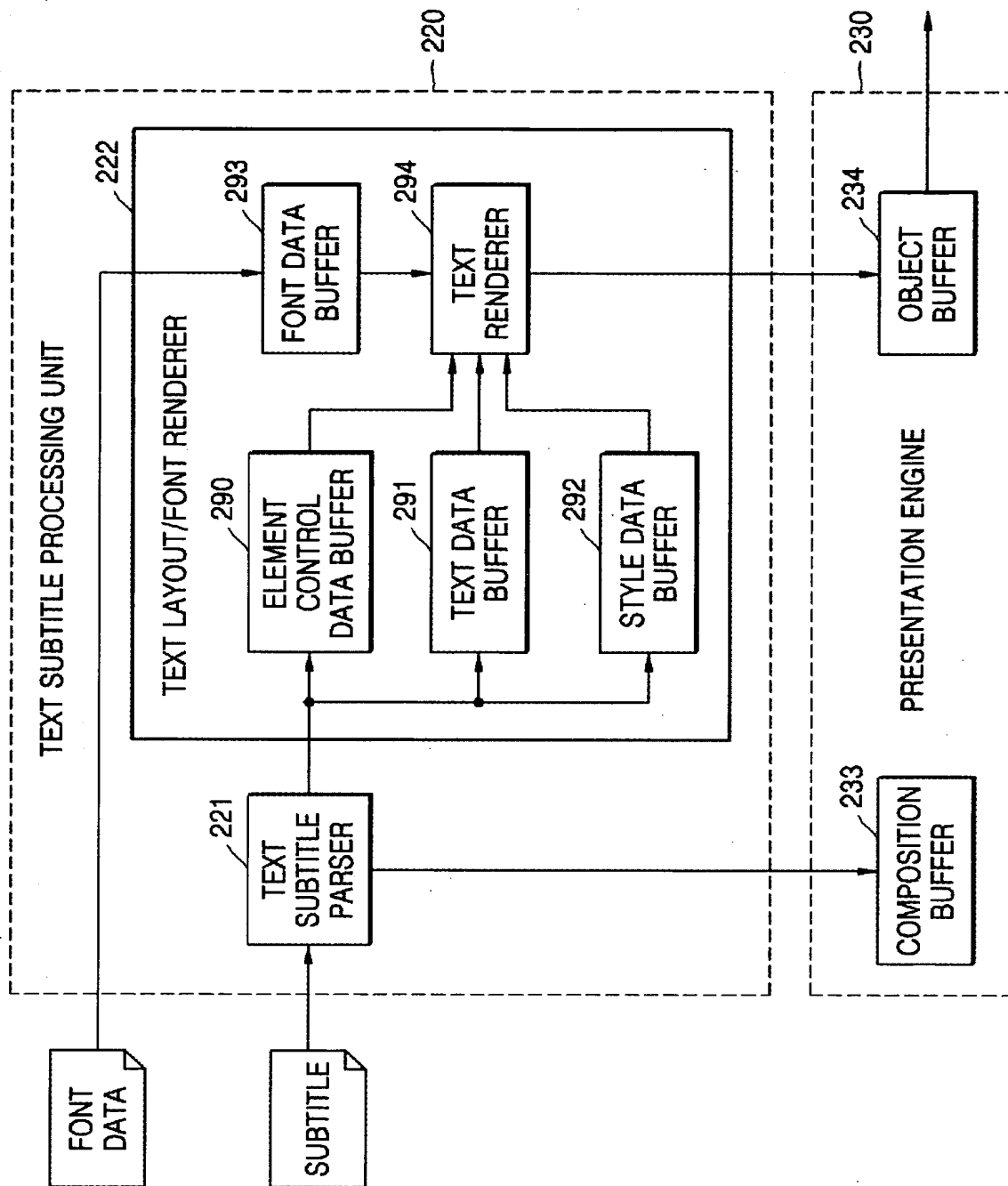


FIG. 4

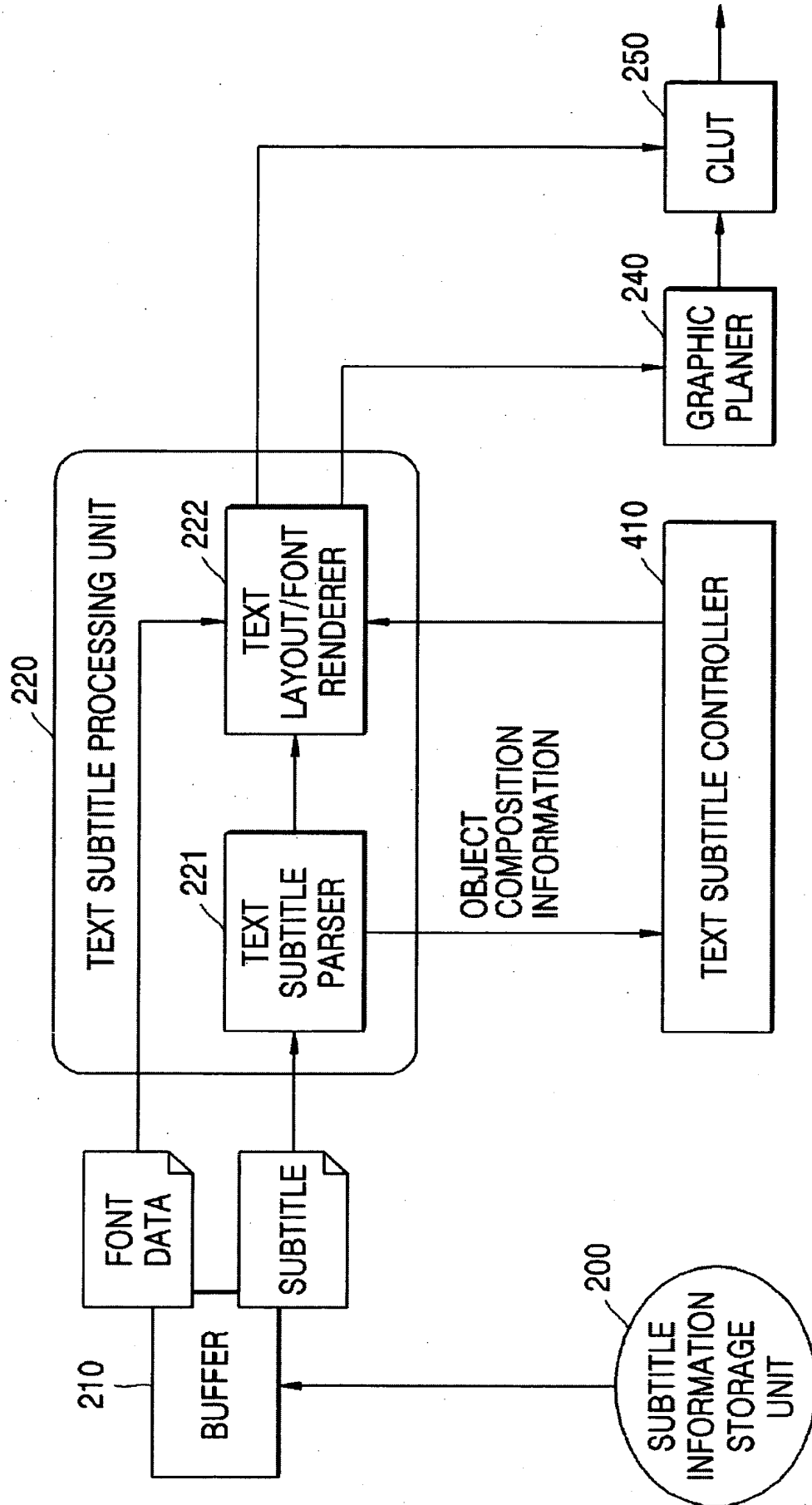


FIG. 5

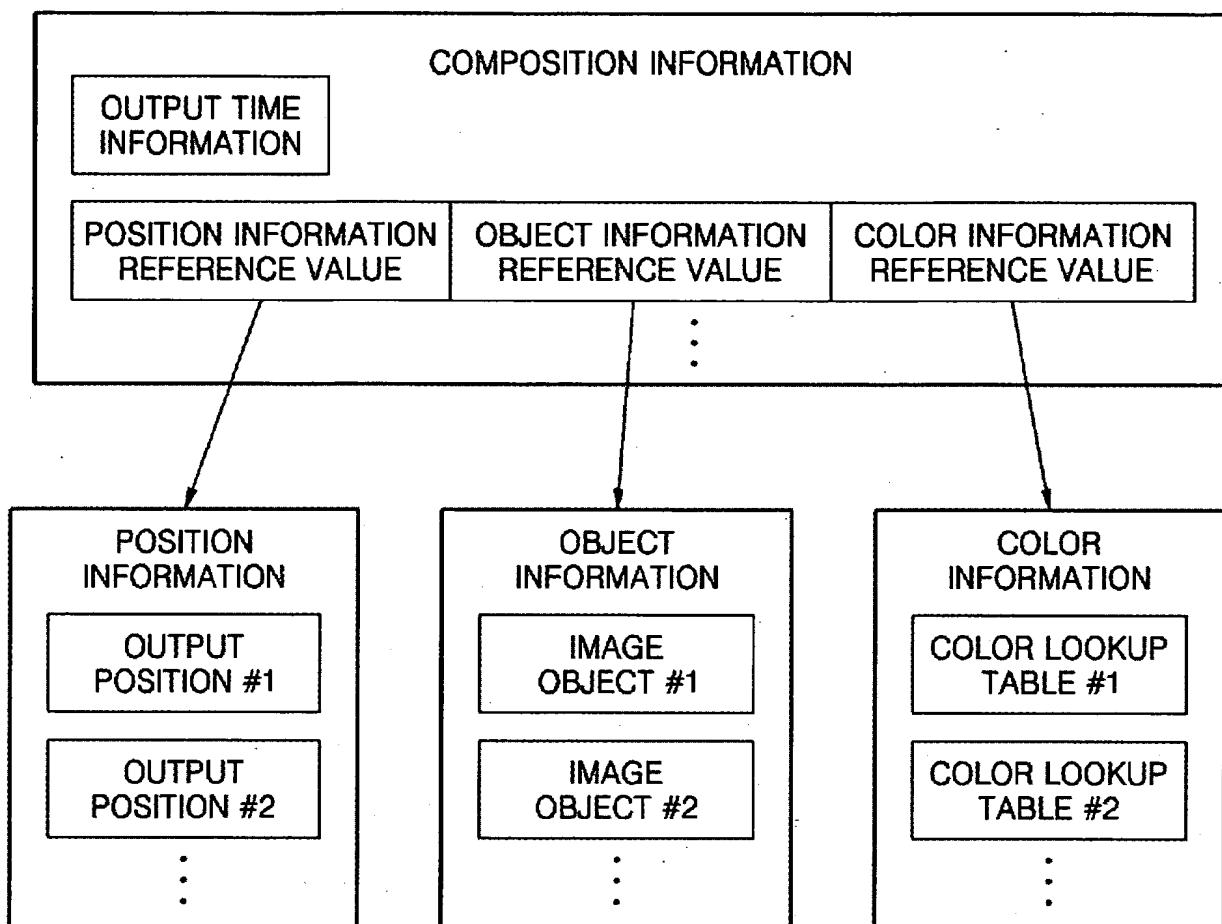


FIG. 6A

```
<style name="Script" font="Arial.ttf" color="black"
background="white" size="16pt" position="x, y" align="center"
direction="left-to-right-top-to-bottom" region="left, top, width,
height" line-height="40px" />
```

FIG. 6B

```
<text start="00:10:00" end="00:15:00" style="Script"> } 610
    Hello
</text>
<text start="00:12:00" end="00:17:00" style="Script"> } 620
    <br/>Subtitle
</text>
<text start="00:14:00" end="00:19:00" style="Script"> } 630
    <br/><br/>World
</text>
```

FIG. 6C

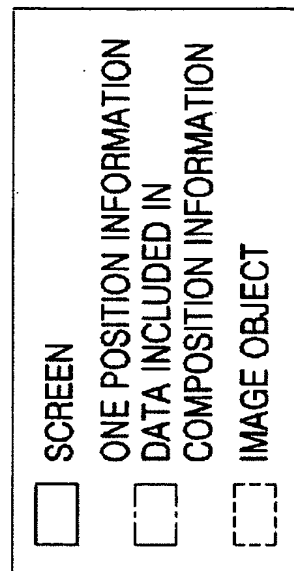
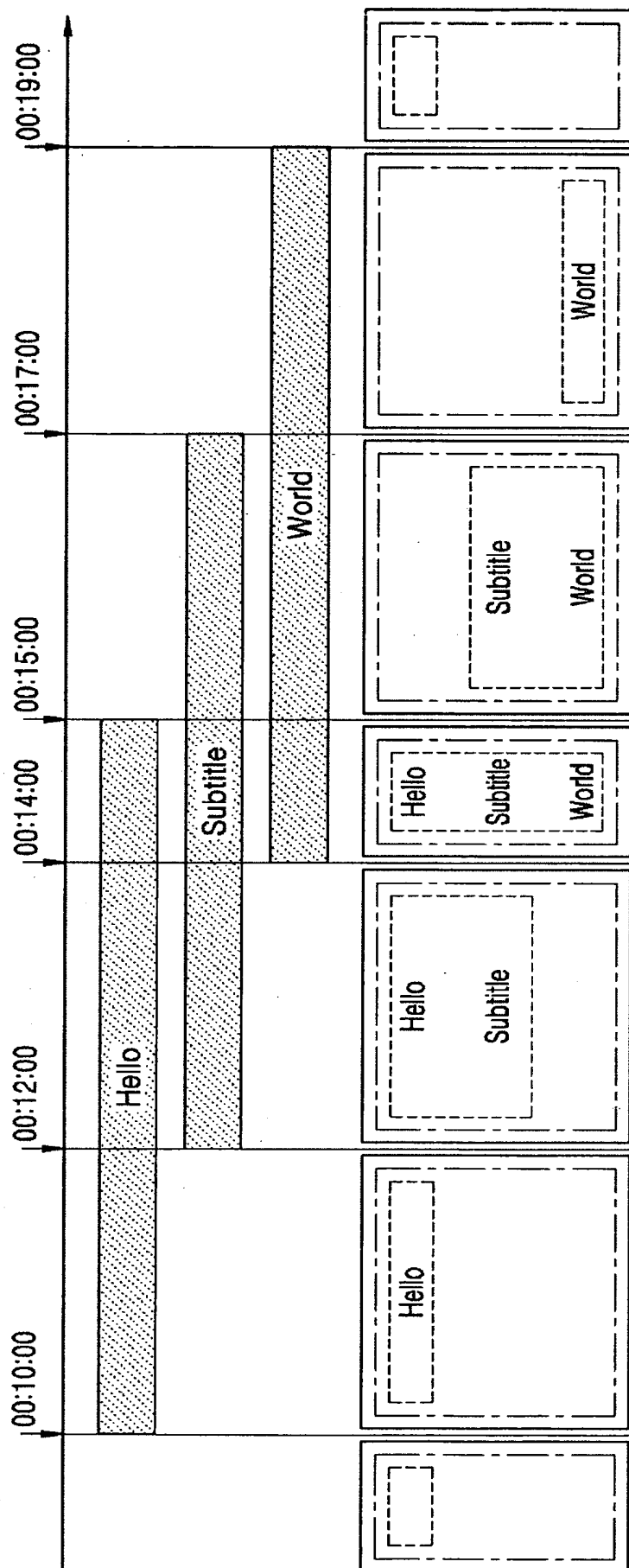


FIG. 7A

```
<style name="Script1" font="Arial.ttf" color="black"
background="white" size="16pt" position="x1, y1" align="center"
direction="left-to-right-top-to-bottom" region="left1, top1, width1,
height1" line-height="40px" />
```

```
<style name="Script2" font="Arial.ttf" color="black"
background="white" size="16pt" position="x2, y2" align="center"
direction="left-to-right-top-to-bottom" region="left2, top2, width2,
height2" line-height="40px" />
```

```
<style name="Script3" font="Arial.ttf" color="black"
background="white" size="16pt" position="x3, y3" align="center"
direction="left-to-right-top-to-bottom" region="left3, top3, width3,
height3" line-height="40px" />
```

FIG. 7B

```
<text start="00:10:00" end="00:15:00" style="Script1">
    Hello } 710
</text>
<text start="00:12:00" end="00:17:00" style="Script2">
    Subtitle } 720
</text>
<text start="00:14:00" end="00:19:00" style="Script3">
    World } 730
</text>
```

FIG. 7C

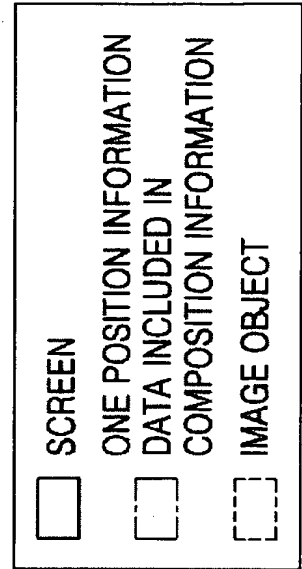
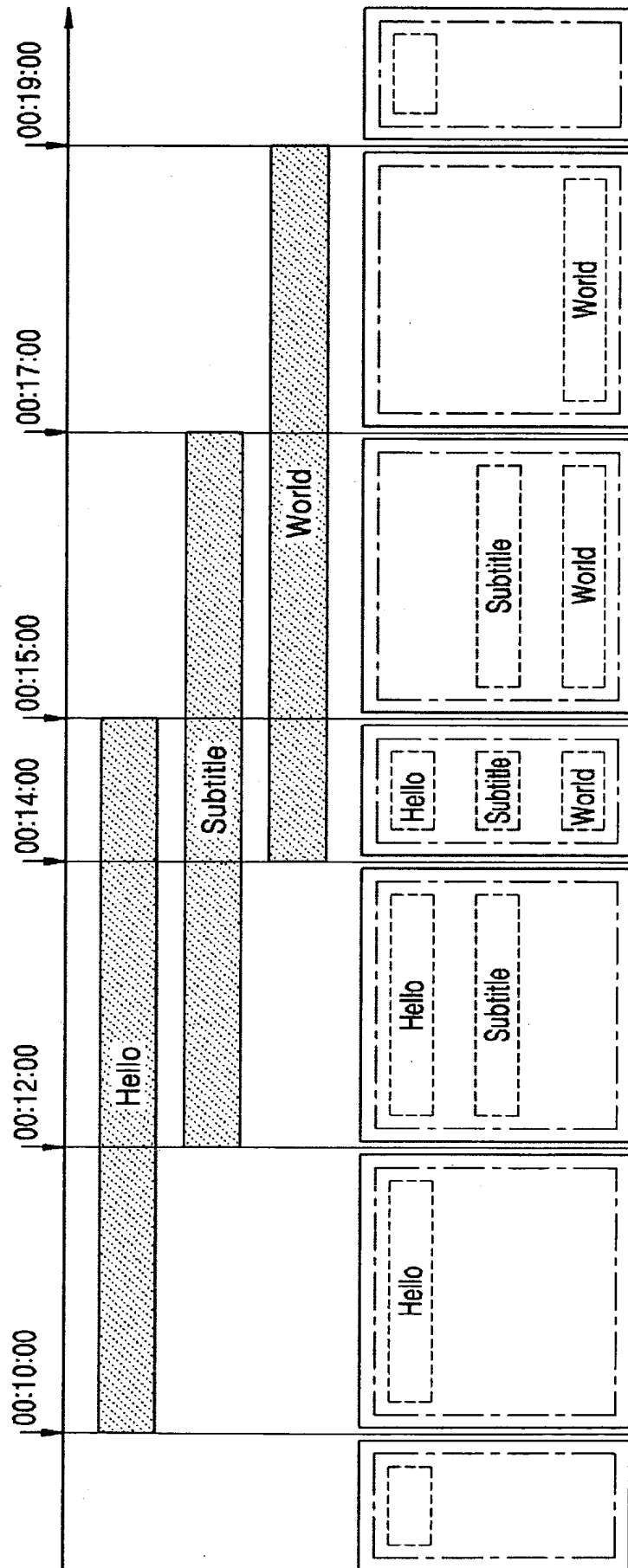


FIG. 8A

```
<style name="Script1" font="Arial.ttf" color="black"
background="white" size="16pt" position="x1, y1" align="center"
direction="left-to-right-top-to-bottom" region="left1, top1, width1,
height1" line-height="40px" />
```

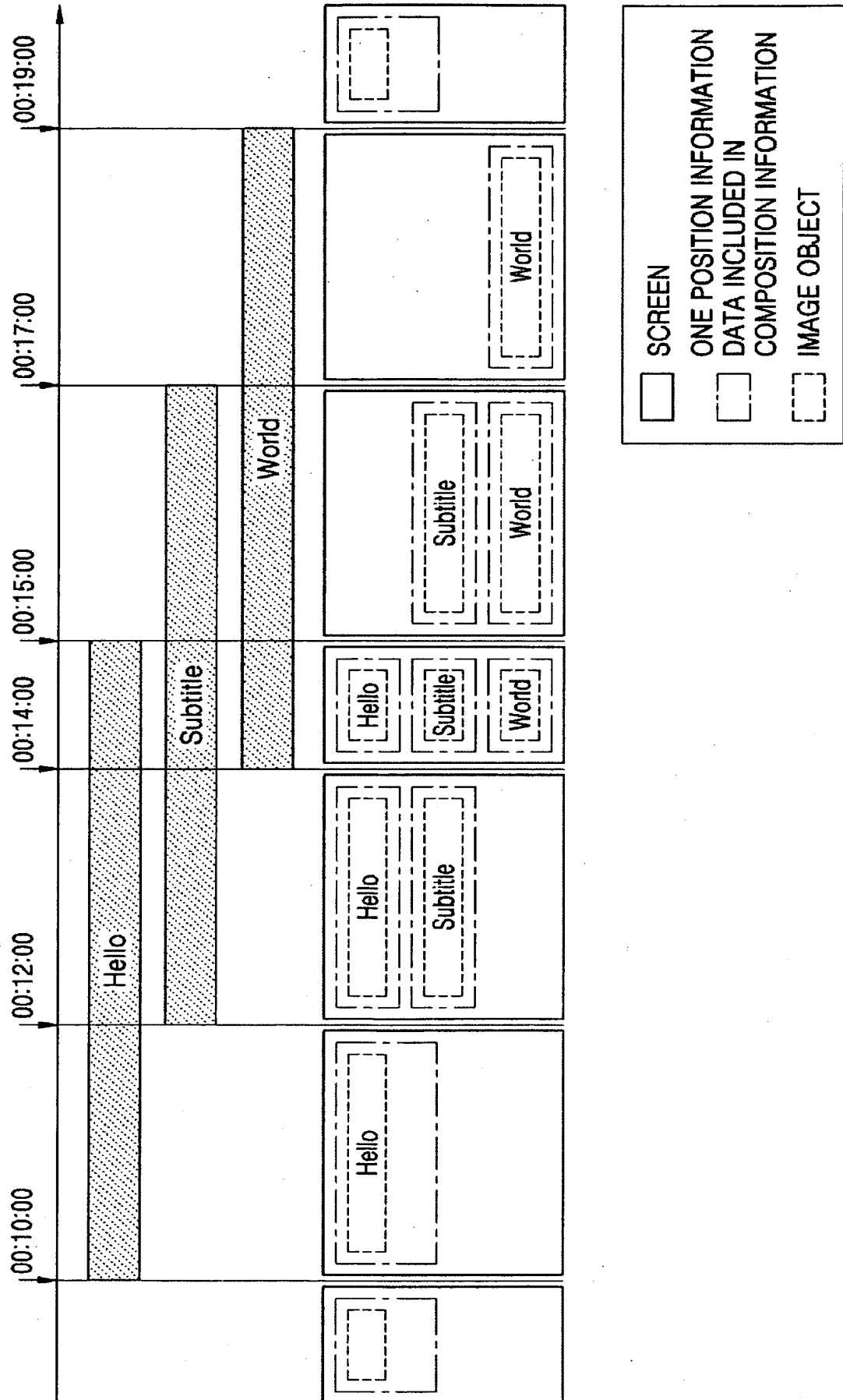
```
<style name="Script2" font="Arial.ttf" color="black"
background="white" size="16pt" position="x2, y2" align="center"
direction="left-to-right-top-to-bottom" region="left2, top2, width2,
height2" line-height="40px" />
```

```
<style name="Script3" font="Arial.ttf" color="black"
background="white" size="16pt" position="x3, y3" align="center"
direction="left-to-right-top-to-bottom" region="left3, top3, width3,
height3" line-height="40px" />
```

FIG. 8B

```
<text start="00:10:00" end="00:15:00" style="Script1"> } 810
    Hello
</text>
<text start="00:12:00" end="00:17:00" style="Script2"> } 820
    Subtitle
</text>
<text start="00:14:00" end="00:19:00" style="Script3"> } 830
    World
</text>
```

FIG. 8C



INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2004/002904

A. CLASSIFICATION OF SUBJECT MATTER**IPC7 G11B 20/10**

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G11B 20/10 G11B 20/12 G11B 27/00 G11B 7/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean Patents and applications for inventions since 1975

Korean utility models and applications for utility models since 1975

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, PAJ "subtitle, text, style, AV(audio-visual), storage, medium, parser, render*, layout"

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y A	EP 0,791,925 A2 (SONY CORPORATION) 27 August 1997 See the whole document	1 7, 17, 21-23
Y A	JP 2001-135066 A (NIPPON COLUMBIA CO., LTD.) 18 May 2001 See the whole document	1 7, 17, 21-23
Y A	JP 2000-322875 A (TOSHIBA CORP.) 24 November 2000 See the whole document	1 7, 17, 21-23
A	WO 92-00647 A1 (WNN VENTURE INC.) 9 January 1992 See the whole document	1, 7, 17, 21-23
P,A	KR 2004-0051446 A (LG ELECTRONICS INC.) 18 June 2004 See the whole document	1, 7, 17, 21-23

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

18 FEBRUARY 2005 (18.02.2005)

Date of mailing of the international search report

23 FEBRUARY 2005 (23.02.2005)

Name and mailing address of the ISA/KR



Korean Intellectual Property Office
920 Dunsan-dong, Seo-gu, Daejeon 302-701,
Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

KIM, Yong Woong

Telephone No. 82-42-481-5698

